Quantum parameter estimation with general dynamics

Haidong Yuan[*]

Department of Mechanical and Automation Engineering,
The Chinese University of Hong Kong, Shatin, Hong Kong

Chi-Hang Fred Fung[†]

Canada Research Centre, Huawei Technologies Canada, Ontario, Canada

One of the main quests in quantum metrology, and quantum parameter estimation in general, is to find out the highest achievable precision with given resources and design schemes that attain that precision. In this article we present a general framework for quantum parameter estimation which relates the ultimate precision limit directly to the geometrical properties of underlying dynamics. With this framework we present systematical methods for computing the ultimate precision limit and optimal probe states. We further demonstrate the power of the framework by deriving a sufficient condition on when ancillary systems are not useful for improving the precision limit.

I. INTRODUCTION

An important task in science and technology is to find the highest achievable precision in measuring and estimating parameters of interest with given resources, and design schemes to reach it. Quantum metrology, which exploits quantum mechanical effects to achieve high precision, has gained increasing attention in recent years[1–19], where a typical situation is to estimate the value of a continuous parameter x encoded in some quantum state ρ_x of the system. To estimate the value, one needs to first perform measurements on the system, which, in the general form, are described by Positive Operator Valued Measurements(POVM), $\{E_y\}$, which provides a distribution for the measurement results $p(y|x) = Tr(E_y\rho_x)$. According to the Cramér-Rao bound in statistical theory[20–23], the standard deviation for any unbiased estimator of x, based on the measurement results y, is bounded below by the Fisher information: $\delta\hat{x} \geq \frac{1}{\sqrt{I(x)}}$, where $\delta\hat{x}$ is the standard deviation of the estimation of x, and $I(x)$ is the Fisher information of the measurement results, $I(x) = \sum_y p(y|x)(\frac{\partial lnp(y|x)}{\partial x})^2$[24]. The Fisher information can be further optimized over all POVMs, which gives

$$\delta\hat{x} \geq \frac{1}{\sqrt{\max_{E_y} I(x)}} = \frac{1}{\sqrt{J(\rho_x)}}, \qquad (1)$$

where the optimized value $J(\rho_x)$ is called quantum Fisher information[22, 23, 25, 26]. If the above process is repeated n times, then the standard deviation of the estimator is bounded by $\delta\hat{x} \geq \frac{1}{\sqrt{nJ(\rho_x)}}$.

To achieve the best precision, we can further optimize the encoding procedures $x \rightarrow \rho_x$ so that $J(\rho_x)$ is maximized. Typically the encoding is achieved by preparing the probe in some initial state ρ_0, then let it evolve under a dynamics which contains the interested parameter, $\rho_0 \xrightarrow{\phi_x} \rho_x$. Usually ϕ_x is determined by a given physical dynamics which is then fixed, while the initial state is up to our choice and can be optimized. A pivotal task in quantum metrology is to find out the optimal initial state ρ_0 and the corresponding maximum quantum Fisher information under any given evolution ϕ_x. When ϕ_x is unitary the GHZ-type of states are known to be optimal which leads to the Heisenberg limit. However when ϕ_x is noisy, such states are in general no longer optimal. Finding the optimal probe states and the corresponding highest precision limit under general dynamics has been the main quest of the field. Recently using the purification approach much progress has been made on developing systematical methods of calculating the highest precision limit[7–10, 13, 14], however how to actually achieve the highest precision limit is still largely unknown, as these methods do not provide ways to obtain the optimal probe states. Another restriction of these methods[7, 8] is that they usually restrict to smooth representations of the Kraus operators, which is not intrinsic to the dynamics.

In this article, we develop a general framework for quantum parameter estimation which relates the ultimate precision limit directly to the geometrical properties of underlying dynamics, this provides systematical methods for computing the ultimate precision limit and optimal probe states without additional assumptions. This framework also provides analytical formulas for the precision limit with arbitrary pure probe states which spares the needs of optimization over equivalent Kraus operators required in previous studies[7, 8]. We further demonstrate the power of the framework by deriving sufficient conditions on when ancillary systems are not useful for improving the precision limit.

---

[*]Electronic address: hdyuan@mae.cuhk.edu.hk
[†]Electronic address: chffung.app@gmail.com

II. ULTIMATE PRECISION LIMIT

The precision limit of measuring x from a set of quantum states ρ_x is determined by the distinguishability between ρ_x and its neighboring states ρ_{x+dx}[25, 30]. This is best seen if we expand the Bures distance between the neighboring states ρ_x and ρ_{x+dx} up to the second order of dx[25]:

$$d_{Bures}^2(\rho_x, \rho_{x+dx}) = \frac{1}{4}J(\rho_x)dx^2, \qquad (2)$$

where $d_{Bures}(\rho_1, \rho_2) = \sqrt{2 - 2F_B(\rho_1, \rho_2)}$, here $F_B(\rho_1, \rho_2) = Tr\sqrt{\rho_1^{\frac{1}{2}}\rho_2\rho_1^{\frac{1}{2}}}$ is the fidelity between two states. Thus maximizing the quantum Fisher information is equivalent as maximizing the Bures distance, which is equivalent as minimizing the fidelity between ρ_x and ρ_{x+dx}. If the evolution is given by ϕ_x, $\rho_x = \phi_x(\rho)$ and $\rho_{x+dx} = \phi_{x+dx}(\rho)$, the problem is then equivalent to finding out $\min_\rho F_B[\phi_x(\rho), \phi_{x+dx}(\rho)]$ and the optimal ρ that achieves the minimum. We now develop tools to solve this problem for both unitary and non-unitary dynamics.

Given two general evolution ϕ_1 and ϕ_2 of the same dimension, we define the Bures angle between them as $B(\phi_1, \phi_2) = \max_\rho \cos^{-1}[F_B(\phi_1(\rho), \phi_2(\rho))]$, this generalizes the Bures angle on quantum states[36]. From the definition of the Bures distance it is easy to see $\max_\rho d_{Bures}^2[\phi_x(\rho), \phi_{x+dx}(\rho)] = 2 - 2\cos B(\phi_x, \phi_{x+dx})$, thus from Eq.(2) we have

$$\max_\rho J[\phi_x(\rho)] = \lim_{dx\to 0} \frac{8[1 - \cos B(\phi_x, \phi_{x+dx})]}{dx^2}. \qquad (3)$$

The ultimate precision limit under the evolution ϕ_x is thus determined by the Bures angle between ϕ_x and the neighboring channels

$$\delta\hat{x} \geq \frac{1}{\lim_{dx\to 0}\frac{\sqrt{8[1-\cos B(\phi_x,\phi_{x+dx})]}}{|dx|}\sqrt{n}}, \qquad (4)$$

where n is the number of times that the procedure is repeated. If ϕ_x is continuous with respect to x, then when $dx \to 0$, $B(\phi_x, \phi_{x+dx}) \to B(\phi_x, \phi_x) = 0$, in this case

$$\begin{aligned}
\max_\rho J[\phi_x(\rho)] &= \lim_{dx\to 0} \frac{8[1 - \cos B(\phi_x, \phi_{x+dx})]}{dx^2} \\
&= \lim_{dx\to 0} \frac{16\sin^2\frac{B(\phi_x,\phi_{x+dx})}{2}}{dx^2} \\
&= \lim_{dx\to 0} \frac{4B^2(\phi_x, \phi_{x+dx})}{dx^2},
\end{aligned} \qquad (5)$$

the ultimate precision limit is then given by

$$\delta\hat{x} \geq \frac{1}{\lim_{dx\to 0}2\frac{B(\phi_x,\phi_{x+dx})}{|dx|}\sqrt{n}}. \qquad (6)$$

The problem is thus reduced to determine the Bures angle between quantum channels. We will first show how to compute the Bures angle between unitary channels, then generalize to noisy quantum channels.

A. Ultimate precision limit for unitary channels

Given two unitaries U_1 and U_2 of the same dimension, since $F_B(U_1\rho U_1^\dagger, U_2\rho U_2^\dagger) = F_B(\rho, U_1^\dagger U_2\rho U_2^\dagger U_1)$, we have $B(U_1, U_2) = B(I, U_1^\dagger U_2)$, i.e., the Bures angle between two unitaries can be reduced to the Bures angle between Identity and a unitary. For a $m \times m$ unitary matrix U, let $e^{-i\theta_j}$ be the eigenvalues of U, where $\theta_j \in (-\pi, \pi]$ for $1 \leq j \leq m$, which we will call the eigen-angles of U. If $\theta_{\max} = \theta_1 \geq \theta_2 \geq \cdots \geq \theta_m = \theta_{\min}$ are arranged in decreasing order, then $B(I, U) = \frac{\theta_{\max}-\theta_{\min}}{2}$ when $\theta_{\max} - \theta_{\min} \leq \pi$[31–33, 39–41], specifically if $U = e^{-iHt}$, then $B(I, U) = \frac{(\lambda_{\max}-\lambda_{\min})t}{2}$ if $(\lambda_{\max} - \lambda_{\min})t \leq \pi$, where $\lambda_{\max(\min)}$ is the maximal (minimal) eigenvalue of H. This provides ways to compute Bures angles on unitary channels. For example, suppose the evolution takes the form $U(x) = (e^{-ixHt})^{\otimes N}$ (tensor product of e^{-ixHt} for N times, which means the same unitary evolution e^{-ixHt} acts on all N probes). Then

$$\begin{aligned}
B[U(x), U(x+dx)] &= B[I, U^\dagger(x)U(x+dx)] \\
&= B[I, (e^{-iHtdx})^{\otimes N}].
\end{aligned} \qquad (7)$$

It is easy to see that the difference between the maximal eigen-angle and the minimal eigen-angle of $(e^{-iHtdx})^{\otimes N}$ is $\theta_{\max} - \theta_{\min} = N(\lambda_{\max} \mid dx \mid t - \lambda_{\min} \mid dx \mid t)$. Thus $B(I, (e^{-iHtdx})^{\otimes N}) = \frac{\theta_{\max}-\theta_{\min}}{2} = \frac{(N\lambda_{\max}|dx|-N\lambda_{\min}|dx|)t}{2}$, Eq.(6) then recovers the Heisenberg limit

$$\delta\hat{x} \geq \frac{1}{\sqrt{n}(\lambda_{\max}-\lambda_{\min})t}\frac{1}{N}.$$

B. Ultimate precision limit for noisy quantum channels

For a general quantum channel which maps from a m_1- to m_2-dimensional Hilbert space, the evolution can be represented by a Kraus operation $K(\rho^S) = \sum_{j=1}^d F_j\rho^S F_j^\dagger$, here the Kraus operators $F_j, 1 \leq j \leq d$, are of the size $m_2 \times m_1$, $\sum_{j=1}^d F_j^\dagger F_j = I_{m_1}$. The channel can be equivalently represented as

$$K(\rho^S) = Tr_E(U_{ES}(|0_E\rangle\langle 0_E| \otimes \rho^S)U_{ES}^\dagger), \qquad (8)$$

where $|0_E\rangle$ denotes some standard state of the environment, and U_{ES} is a unitary operator acting on both system and environment, which we will call as the unitary

extension of K. A general U_{ES} can be written as

$$U_{ES} = (W_E \otimes I_{m_2}) \underbrace{\begin{bmatrix} F_1 & * & * & \cdots & * \\ F_2 & * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ F_d & * & * & \cdots & * \\ \mathbf{0} & * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & * & * & \cdots & * \end{bmatrix}}_{U}, \qquad (9)$$

here only the first m_1 columns of U are fixed and $W_E \in U(p)(p \times p$ unitaries) only acts on the environment and can be chosen arbitrarily, here $p \geq d$ as $p - d$ zero Kraus operators can be added.

Given a channel an ancillary system can be used to improve the precision limit, this can be described as the extended channel

$$(K \otimes I_A)(\rho^{SA}) = \sum_j (F_j \otimes I_A)\rho^{SA}(F_j \otimes I_A)^\dagger,$$

here ρ^{SA} represents a state of the original and ancillary systems. Without loss of generality, the ancillary system can be assumed to have the same dimension as the original system.

Given two quantum channels K_1 and K_2 of the same dimension, let U_{ES1} and U_{ES2} as unitary extensions of K_1 and K_2 respectively, we have[34]

$$\begin{aligned} B(K_1 \otimes I_A, K_2 \otimes I_A) &= \min_{U_{ES1},U_{ES2}} B(U_{ES1}, U_{ES2}) \\ &= \min_{U_{ES1}} B(U_{ES1}, U_{ES2}) \qquad (10) \\ &= \min_{U_{ES2}} B(U_{ES1}, U_{ES2}). \end{aligned}$$

This extends Uhlmann's purication theorem on mixed states[48] to noisy quantum channels. Furthermore we show in the appendix that $B(K_1 \otimes I_A, K_2 \otimes I_A)$ can be explicitly computed from the Kraus operators of K_1 and K_2[38]: if $K_1(\rho^S) = \sum_{j=1}^d F_{1j}\rho^S F_{1j}^\dagger$, $K_2(\rho^S) = \sum_{j=1}^d F_{2j}\rho^S F_{2j}^\dagger$, then $\cos B(K_1 \otimes I_A, K_2 \otimes I_A) = \max_{\|W\| \leq 1} \frac{1}{2}\lambda_{\min}(K_W + K_W^\dagger)$, here $\lambda_{\min}(K_W + K_W^\dagger)$ denotes the minimum eigenvalue of $K_W + K_W^\dagger$ where $K_W = \sum_{ij} w_{ij} F_{1i}^\dagger F_{2j}$, with w_{ij} as the ij-th entry of a $d \times d$ matrix W which satisfies $\|W\| \leq 1(\|\cdot\|$ denotes the operator norm which equals to the maximum singular value). If we substitute $K_1 = K_x$ and $K_2 = K_{x+dx}$, where $K_x(\rho^S) = \sum_{j=1}^d F_j(x)\rho^S F_j^\dagger(x)$ and $K_{x+dx}(\rho^S) = \sum_{j=1}^d F_j(x+dx)\rho^S F_j^\dagger(x+dx)$ with x being the interested parameter, then

$$\begin{aligned} &\cos B(K_x \otimes I_A, K_{x+dx} \otimes I_A) \\ &= \max_{\|W\| \leq 1} \frac{1}{2}\lambda_{\min}(K_W + K_W^\dagger), \end{aligned} \qquad (11)$$

where $K_W = \sum_{ij} w_{ij} F_i^\dagger(x)F_j(x + dx)$. By substituting $\phi_x = K_x \otimes I_A$ and $\phi_{x+dx} = K_{x+dx} \otimes I_A$ in Eq.(3), we then get the maximal quantum Fisher information for the extended channel $K_x \otimes I_A$,

$$\max J = \lim_{dx\to 0} \frac{8[1 - \max_{\|W\| \leq 1} \frac{1}{2}\lambda_{\min}(K_W + K_W^\dagger)]}{dx^2}.(12)$$

In previous studies the operator W_E in Eq.(9), which can be arbitrary chosen, was assumed to depend on x smoothly[7, 8]. As a result, the W in Eq.(12) was restricted to unitary operators that depends smoothly on x as explained in detail in the appendix A 2. This restriction was introduced out of computational convenience in previous studies, which is not intrinsic to the dynamics. The formula here does not have such assumption and can be applied more broadly, for example it can be applied to the discrimination of quantum channels which is discrete in nature[34]. Also since any W that is not optimal gives a lower bound on the precision limit, the formula here also provides more room for obtaining useful lower bounds.

The maximization in Eq.(12) can be further formulated as a semi-definite programming and solved efficiently: $\max_{\|W\| \leq 1} \frac{1}{2}\lambda_{\min}(K_W + K_W^\dagger) =$

$$\begin{aligned} maximize \quad & \frac{1}{2}t \\ s.t. \quad & \begin{pmatrix} I & W^\dagger \\ W & I \end{pmatrix} \succeq 0, \qquad (13) \\ & K_W + K_W^\dagger - tI \succeq 0. \end{aligned}$$

Another advantage of this formulation is that the dual form of this semi-definite programming provides a systematical way for obtaining the optimal probe states, which we will show in the next setion.

III. OPTIMAL PROBE STATES

Developing systematical methods to obtain the optimal probe states are essential for achieving the precision limit. So far there are only a few cases for which optimal probe states are known, mostly for phase estimations[11, 12, 42–45]. A systematical way of obtaining optimal probe states for general quantum dynamics is highly desired as it will pave the way for achieving the ultimate precision limit. We now show how to obtain the optimal probe states that achieve the ultimate precision limit for extended channels.

We first provide an analytical formula for calculating quantum Fisher information with any given pure input states, for both unextended and extended channels, then use it to obtain optimal probe states for extended channels.

In the appendix we showed that for both unextended and extended channels with pure probe states we have[38]

$$F_B[K_x(\rho^S), K_{x+dx}(\rho^S)] = \|M(\rho^S)\|_1, \qquad (14)$$

$$F_B[(K_x \otimes I_A)(\rho^{SA}), (K_{x+dx} \otimes I_A)(\rho^{SA})] = \|M(\rho^S)\|_1,$$
$$(15)$$

here $M(\rho^S)$ is a $d \times d$ matrix with its ij-entry equals to $Tr[\rho^S F_i^\dagger(x) F_j(x+dx)]$, and $\|\cdot\|_1$ represents the trace norm which equals to the summation of singular values. For the unextended channel this formula works for the pure probe state $\rho^S = |\psi_S\rangle\langle\psi_S|$, while for the extended channel although $\rho^{SA} = |\psi_{SA}\rangle\langle\psi_{SA}|$ is required to be a pure state, $\rho^S = Tr_A(\rho^{SA})$ can be any mixed state, which characterizes the advantage provided by ancillary systems.

The above two formulas provide a straightforward way calculating the quantum Fisher information with any pure probe states,

$$J[K_x(\rho^S)] = \lim_{dx \to 0} \frac{8(1 - \|M(\rho^S)\|_1)}{dx^2},$$
$$J[(K_x \otimes I_A)(\rho^{SA})] = \lim_{dx \to 0} \frac{8(1 - \|M(\rho^S)\|_1)}{dx^2}. \quad (16)$$

In contrast to previous studies[7, 8, 28], optimization over equivalent representations of Kraus operator is not needed in this formulation. In fact $\|M(\rho^S)\|_1$ does not depend on any particular representation of the Kraus operators: if we use a different representation of Kraus operators for K_x and K_{x+dx}, for example $\tilde{F}_i(x) = \sum_r u_{ir} F_r(x)$ and $\tilde{F}_j(x+dx) = \sum_s v_{js} F_s(x+dx)$, where u_{ir} and v_{js} are entries of some unitary matrices U and V respectively, then $Tr[\rho^S \tilde{F}_i^\dagger(x) \tilde{F}_j(x+dx)] = \sum_{r,s} u_{ir}^* v_{js} Tr[\rho^S F_r^\dagger(x) F_s(x+dx)]$, thus $\tilde{M}(\rho^S) = \bar{U} M(\rho^S) V^T$ which has the same trace norm $\|\tilde{M}(\rho^S)\|_1 = \|M(\rho^S)\|_1$.

The optimal probe states for the extended channel can then be obtained by minimizing over input states at both sides of Eq.(15),

$$\min_{\rho^{SA}} F_B[(K_x \otimes I_A)(\rho^{SA}), (K_{x+dx} \otimes I_A)(\rho^{SA})]$$
$$= \min_{\rho^S} \|M(\rho^S)\|_1. \quad (17)$$

This can be computed by a semi-definite programming formulation for the trace norm[46] as $\min_{\rho^S} \|M(\rho^S)\|_1 =$

$$\begin{aligned} minimize \quad & \frac{1}{2} Tr(P) + \frac{1}{2} Tr(Q) \\ s.t. \quad & \begin{pmatrix} P & M^\dagger(\rho^S) \\ M(\rho^S) & Q \end{pmatrix} \succeq 0, \quad (18) \\ & \rho^S \succeq 0, Tr(\rho^S) = 1, \end{aligned}$$

where P, Q are Hermitian matrices. One can verify that this is exactly the dual form of the semi-definite programming used in Eq.(13). From the output ρ^S, we can easily obtain the optimal probe state ρ^{SA}, which is any purification with the reduced states equals to the optimal ρ^S. This gives a systematical way to obtain the optimal probe states for the extended channel which we demonstrate through some examples.

Consider phase estimation with spontaneous emission, $K_x(\rho_0) = F_1(x)\rho_0 F_1^\dagger(x) + F_2(x)\rho_0 F_2^\dagger(x)$, where $F_1(x) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{\eta} \end{pmatrix} U(x), F_2(x) = \begin{pmatrix} 0 & \sqrt{1-\eta} \\ 0 & 0 \end{pmatrix} U(x)$, $U(x) = \exp(-i\frac{\sigma_3}{2}x)$. Suppose a pure input state ρ^{SA} is prepared for the extended channel $K_x \otimes I_A$, then $\rho^S = Tr_A(\rho^{SA})$, $M(\rho^S) = \begin{pmatrix} Tr[\rho^S F_1^\dagger(x) F_1(x+dx)] & Tr[\rho^S F_1^\dagger(x) F_2(x+dx)] \\ Tr[\rho^S F_2^\dagger(x) F_1(x+dx)] & Tr[\rho^S F_2^\dagger(x) F_2(x+dx)] \end{pmatrix}$, in this case the problem can be solved analytically, in the appendix we showed that the optimal ρ^S is given by $\rho^S = \begin{pmatrix} \frac{\sqrt{\eta}}{1+\sqrt{\eta}} & 0 \\ 0 & \frac{1}{1+\sqrt{\eta}} \end{pmatrix}$ and the corresponding maximal quantum Fisher information is $\max J = \frac{4\eta}{(1+\sqrt{\eta})^2}$[38]. Since the optimal ρ^S is mixed, an ancillary systems is necessary. The optimal input state in this case is any pure state ρ^{SA} with the reduced state equal to ρ^S, the simplest choice of the optimal input state in this case is $\sqrt{\frac{\sqrt{\eta}}{1+\sqrt{\eta}}}|00\rangle + \sqrt{\frac{1}{1+\sqrt{\eta}}}|11\rangle$, which is not a maximally entangled state as previously suspected[13, 37]. We can also use the method to find the maximal quantum Fisher information without using ancilla by imposing the condition that ρ^S be pure. In that case the maximal quantum Fisher information turns out to be η and the optimal input state has the form $(|0\rangle + \exp(i\theta)|1\rangle)/\sqrt{2}$ for some $\theta \in \mathbb{R}$[38].

For high dimensional systems, we use the CVX package in Matlab[47] to implement the semi-definite programming of (18) and obtain the optimal input states. For example, consider two qubits with independent dephasing noises, which can be represented by a Kraus operation with 4-Kraus operators: $F_1(x) \otimes F_1(x), F_1(x) \otimes F_2(x), F_2(x) \otimes F_1(x), F_2(x) \otimes F_1(x)$ with $F_1(x) = \sqrt{\frac{1+\eta}{2}} U(x)$, $F_2(x) = \sqrt{\frac{1-\eta}{2}} \sigma_3 U(x)$, here $U(x) = \exp(-i\frac{\sigma_3}{2}x)$. It turns out that $\min_{\rho^S} \|M(\rho^S)\|$ can always be attained with a pure ρ^S, ancillary systems are thus not necessary. In Fig. 1 we plotted the entanglement of optimal states, which is quantified by the entropy of the reduced single qubit state ρ^S, at different η. It can be seen that there exists a threshold for η: when η exceeds the threshold, the optimal state is the GHZ state, which is maximally entangled; when η is below the threshold, GHZ state ceases to be optimal, with the decreasing of η, the optimal state gradually changes from the maximally entangled state to separable state. Fig. 2 shows the quantum Fisher information with the optimal state and the separable input state $|++\rangle$, where $|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$. It can be seen that the gain of entanglement is only obvious in the region of high η, i.e., low noises. Similar behaviour is found for more qubits, i.e., there exists a threshold for η, above the threshold the optimal state is the GHZ state and with the decreasing of η, the optimal state gradually changes from GHZ state to separable state, and this threshold increases with the number
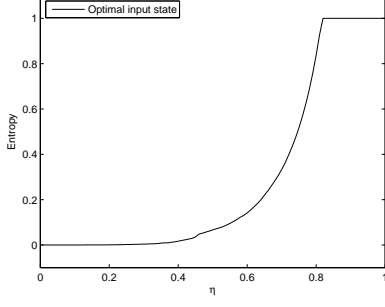
FIG. 1: Entropy of reduced single qubit state, which is used to quantify the entanglement of the optimal state, at different η.
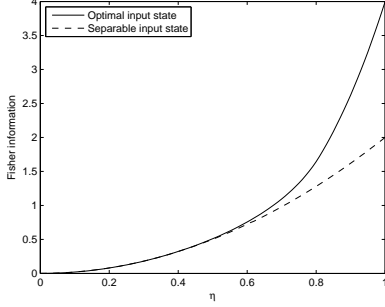

FIG. 2: Quantum Fisher information with the optimal input state and separable input state $|++\rangle$ for 2 qubits with independent dephasing noises.
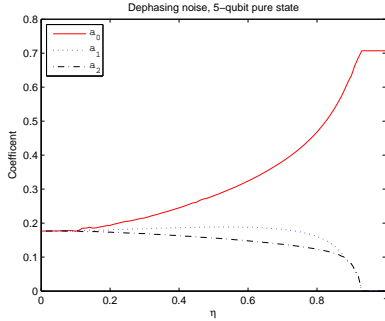

FIG. 3: Optimal probe state for 5 qubits under independent dephasing noises. The optimal state has the form $|\psi\rangle = a_0|\psi_0\rangle + a_1|\psi_1\rangle + a_2|\psi_2\rangle$, where $|\psi_i\rangle$ denotes the summation of all basis states with i zeros or i ones, for example $|\psi_0\rangle = |00000\rangle + |11111\rangle$.

of qubits. In Fig.3 the optimal state for 5 qubits with independent dephasing noises is shown, and in Fig.4 the quantum Fisher information for the optimal state, GHZ state and the separable state are plotted. We also calculated the optimal state for 10 qubits with 5 of them under independent spontaneous emission and 5 of them as ancillary qubits, in Fig.5 we plotted the Fisher information for the optimal state, GHZ state and separable state.
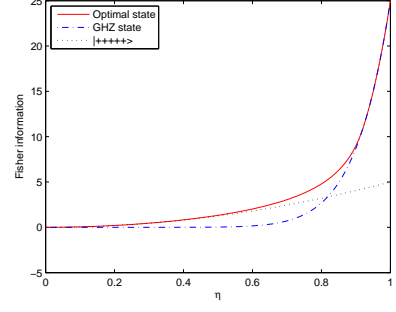

FIG. 4: Quantum Fisher information for optimal probe states, GHZ state and separable state for 5 qubits under independent dephasing noises.
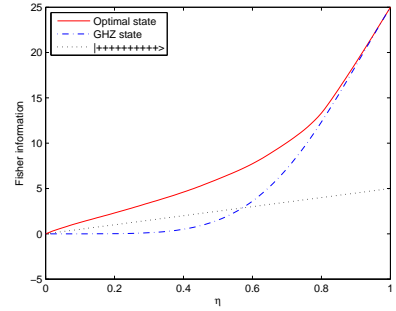

FIG. 5: Quantum Fisher information for optimal probe states, GHZ state and separable state for 10 qubits with 5 of them under independent spontaneous emission and 5 of them as ancillary system..

IV. WHEN ANCILLARY SYSTEMS ARE NOT HELPFUL

The formulas developed here not only provides systematical methods to compute the ultimate precision limit and optimal probe states, but also have wide implications, which we will demonstrate by deriving a sufficient condition on when ancillary systems are not useful for improving the precision limit.

We have seen that in the spontaneous emission case ancillary system helps improving the precision limit while in some other cases—known examples include unitary, classical, phase estimation with dephasing and lossy channels[13, 37], ancillary systems do not help. For a general channel it is usually difficult to tell whether ancillary systems can help improving the precision limit or not. Previously this problem was usually studied case by case by comparing the maximum quantum Fisher information for the unextended and extended channels. The formulas developed here provide a more direct way: from Eq.(16) it is obvious that ancillary systems do not help improving the precision limit if and only if $\min_{\rho^S} \|M(\rho^S)\|_1$ can be reached at a pure state ρ^S, which can be checked by using the semi-definite programming to obtain op-

timal ρ^S and see if $(\rho^S)^2 = \rho^S$. A more easily verifiable sufficient condition is as following: given a channel $K_x(\rho) = \sum_{i=1}^d F_i(x)\rho F_i^\dagger(x)$, if $F_i^\dagger(x)F_j(x + dx)$, $1 \le i, j \le d$ can be simultaneously diagonalized, then ancillary systems do not help improving the precision limit. As if there exist a basis such that $F_i^\dagger(x)F_j(x + dx)$ are all diagonal, then in that basis only the diagonal entries of ρ^S enter into $M(\rho^S)$(as the entries of $M(\rho^S)$ are of the form $Tr[\rho^S F_i^\dagger(x)F_j(x + dx)]$ which only depends on the diagonal terms of ρ^S if $F_i^\dagger(x)F_j(x + dx)$ is diagonal), other entries can be chosen freely. This means that $\min_{\rho^S} \|M(\rho^S)\|_1$ can be achieved by all ρ^S with the optimal diagonal entries, $\rho_{ii}^S = a_i$, this always includes a pure state $|\psi_S\rangle = \sum_i \sqrt{a_i}|i\rangle$, hence ancillary system is not necessary to achieve $\min_{\rho^S} \|M(\rho^S)\|_1$ for such channels. This sufficient condition is satisfied by unitary, classical and phase estimation with dephasing channel, and many other channels that have not been categorized before, for example phase estimation with noises along the X and Y directions satisfies this condition, as it can be represented with two following Kraus operators $F_1(x) = \sqrt{\frac{1+\eta}{2}}\sigma_1 \exp(-i\frac{\sigma_3}{2}x)$ and $F_2(x) = \sqrt{\frac{1-\eta}{2}}\sigma_2 \exp(-i\frac{\sigma_3}{2}x)$, one can easily check that they satisfy the condition thus ancillary system does not help improving the ultimate precision limit for this channel.

V. CONCLUSION

In conclusion we presented a general framework for quantum metrology which provides systematical ways of obtaining the ultimate precision limit and optimal input states. This framework relates the ultimate precision limit directly to the underlying dynamics, which opens the possibility of utilizing quantum control methods to alter the underlying dynamics for better precision limit[35]. The tools developed here, such as the generalized Bures angle on quantum channels that can be efficiently computed using semi-definite programming, are expected to find wide applications in various fields of quantum information science.

Appendix A: Appendix

In the appendix we give detailed derivation on the Bures angle between extended channels, the analytical formula to compute quantum Fisher information with arbitrary pure input state and two examples with dephasing noise and spontaneous emission respectively.

1. Bures angle between extended channels

In this section we show that for any two given channels which maps from m_1- to m_2-dimensional Hilbert space, $K_1(\rho^S) = \sum_{i=1}^d F_{1i}\rho^S F_{1i}^\dagger$ and $K_2(\rho^S) = \sum_{i=1}^d F_{2i}\rho^S F_{2i}^\dagger$, the Bures angle between the extended channels $K_1 \otimes I_A$

and $K_2 \otimes I_A$ can be computed from the Kraus operators as following

$$\cos B(K_1 \otimes I_A, K_2 \otimes I_A) = \max_{\|W\| \le 1} \frac{1}{2}\lambda_{\min}(K_W + K_W^\dagger),$$

i.e.,

$$\begin{aligned}
&\min_{\rho^{SA}} F_B[(K_1 \otimes I_A)(\rho^{SA}), (K_2 \otimes I_A)(\rho^{SA})] \\
&= \max_{\|W\| \le 1} \frac{1}{2}\lambda_{\min}(K_W + K_W^\dagger),
\end{aligned} \tag{A1}$$

here $K_W = \sum_{ij} w_{ij} F_{1i}^\dagger F_{2j}$, w_{ij} is the ij-th entry of $d \times d$ matrix W and the optimization is over all $d \times d$ matrices W with operator norm $\|W\| \le 1$. During the proof the analytical formulas that compute quantum Fisher information for arbitrary pure input states will be derived.

As the minimum fidelity can always be achieved with pure states, we can assume $\rho^{SA} = |\psi_{SA}\rangle\langle\psi_{SA}|$. Denote U_{ES1}, U_{ES2} as the unitary extension of K_1 and K_2 respectively, i.e.,

$$\begin{aligned}
K_1(\rho^S) &= Tr_E(U_{ES1}(|0_E\rangle\langle 0_E| \otimes \rho^S)U_{ES1}^\dagger), \\
K_2(\rho^S) &= Tr_E(U_{ES2}(|0_E\rangle\langle 0_E| \otimes \rho^S)U_{ES2}^\dagger),
\end{aligned} \tag{A2}$$

where $|0_E\rangle$ denotes some standard state of the environment, U_{ES1} and U_{ES2} are unitary operators acting on both system and environment. The general form of U_{ES1} can be written as a pm_2-dimensional unitary, here $p \ge d$, $p - d$ zero Kraus operators can be added,

$$U_{ES1} = (W_1 \otimes I_{m_2}) \underbrace{\begin{bmatrix} F_{11} & * & * & \cdots & * \\ F_{12} & * & * & \cdots & * \\ \vdots & \vdots & & & \vdots \\ F_{1d} & * & * & \cdots & * \\ \mathbf{0} & * & * & \cdots & * \\ \vdots & \vdots & & & \vdots \\ \mathbf{0} & * & * & \cdots & * \end{bmatrix}}_{U_1} \tag{A3}$$

where $W_1 \in U(p)$ and only the first m_1 columns of U_1 are fixed. Similarly

$$U_{ES2} = (W_2 \otimes I_{m_2}) \underbrace{\begin{bmatrix} F_{21} & * & * & \cdots & * \\ F_{22} & * & * & \cdots & * \\ \vdots & \vdots & & & \vdots \\ F_{2d} & * & * & \cdots & * \\ \mathbf{0} & * & * & \cdots & * \\ \vdots & \vdots & & & \vdots \\ \mathbf{0} & * & * & \cdots & * \end{bmatrix}}_{U_2} \tag{A4}$$

where $W_2 \in U(p)$ and only the first m_1 columns of U_2 are fixed. Note that W_1 and W_2 only act on the environment and can be chosen arbitrarily.

$$U^\dagger_{ES1}U_{ES2} = \begin{bmatrix} F^\dagger_{11} & F^\dagger_{12} & * & \cdots & F^\dagger_{1d} & \mathbf{0} & * & \cdots & \mathbf{0} \\ * & * & * & \cdots & * & * & * & \cdots & * \\ \vdots & & & & \vdots & & & & \vdots \\ * & * & * & \cdots & * & * & * & \cdots & * \end{bmatrix} (W^\dagger_1 \otimes I_{m_2})(W_2 \otimes I_{m_2}) \begin{bmatrix} F_{21} & * & * & \cdots & * \\ F_{22} & * & * & \cdots & * \\ \vdots & & & & \vdots \\ F_{2d} & * & * & \cdots & * \\ \mathbf{0} & * & * & \cdots & * \\ \vdots & & \vdots & & \vdots \\ \mathbf{0} & * & * & \cdots & * \end{bmatrix}$$

$$= \begin{bmatrix} F^\dagger_{11} & F^\dagger_{12} & * & \cdots & F^\dagger_{1d} & \mathbf{0} & * & \cdots & \mathbf{0} \\ * & * & * & \cdots & * & * & * & \cdots & * \\ \vdots & & & & \vdots & & & & \vdots \\ * & * & * & \cdots & * & * & * & \cdots & * \end{bmatrix} (W_p \otimes I_{m_2}) \begin{bmatrix} F_{21} & * & * & \cdots & * \\ F_{22} & * & * & \cdots & * \\ \vdots & & & & \vdots \\ F_{2d} & * & * & \cdots & * \\ \mathbf{0} & * & * & \cdots & * \\ \vdots & & \vdots & & \vdots \\ \mathbf{0} & * & * & \cdots & * \end{bmatrix} \tag{A5}$$

$$= \begin{bmatrix} K_{W_d} & * & * & \cdots & * \\ * & * & * & \cdots & * \\ \vdots & & & & \vdots \\ * & * & * & \cdots & * \end{bmatrix},$$

where $W_p = W^\dagger_1 W_2 \in U(p)$, W_d is the first $d \times d$ submatrix of W_p, i.e.,

$$W_p = \begin{bmatrix} W_d & * \\ * & * \end{bmatrix} \tag{A6}$$

and $K_{W_d} = \sum_{i=1}^{d} \sum_{j=1}^{d} w_{ij} F^\dagger_{1i} F_{2j}$ with w_{ij} as the ij-th entry of W_d. Then

$$\min_{|\psi_{SA}\rangle} F_B[(K_1 \otimes I_A)(|\psi_{SA}\rangle\langle\psi_{SA}|), (K_2 \otimes I_A)(|\psi_{SA}\rangle\langle\psi_{SA}|)]$$

$$= \min_{|\psi_{SA}\rangle} F_B[Tr_E(U_{ES1} \otimes I_A(|0_E\rangle\langle 0_E| \otimes |\psi_{SA}\rangle\langle\psi_{SA}|)U^\dagger_{ES1} \otimes I_A), Tr_E(U_{ES2} \otimes I_A(|0_E\rangle\langle 0_E| \otimes |\psi_{SA}\rangle\langle\psi_{SA}|)U^\dagger_{ES2} \otimes I_A)]$$

$$= \min_{|\psi_{SA}\rangle} \max_{U_{ES2}} F_B(U_{ES1} \otimes I_A(|0_E\rangle\langle 0_E| \otimes |\psi_{SA}\rangle\langle\psi_{SA}|)U^\dagger_{ES1} \otimes I_A, U_{ES2} \otimes I_A(|0_E\rangle\langle 0_E| \otimes |\psi_{SA}\rangle\langle\psi_{SA}|)U^\dagger_{ES2} \otimes I_A),$$

the second equality used Uhlmann's theorem which states that

$$F_B(\rho_1, \rho_2) = \max_{|\psi_2\rangle} F_B(|\psi_1\rangle\langle\psi_1|, |\psi_2\rangle\langle\psi_2|), \tag{A7}$$

where $|\psi_1\rangle$ is an arbitrary purification of ρ_1 and the maximization runs over all purifications $|\psi_2\rangle$ of ρ_2[48]. Thus

$$\min_{|\psi_{SA}\rangle} F_B[(K_1 \otimes I_A)(|\psi_{SA}\rangle\langle\psi_{SA}|), (K_2 \otimes I_A)(|\psi_{SA}\rangle\langle\psi_{SA}|)]$$

$$= \min_{|\psi_{SA}\rangle} \max_{U_{ES2}} F_B[U_{ES1} \otimes I_A(|0_E\rangle\langle0_E| \otimes |\psi_{SA}\rangle\langle\psi_{SA}|)U_{ES1}^\dagger \otimes I_A, U_{ES2} \otimes I_A(|0_E\rangle\langle0_E| \otimes |\psi_{SA}\rangle\langle\psi_{SA}|)U_{ES2}^\dagger \otimes I_A]$$

$$= \min_{|\psi_{SA}\rangle} \max_{U_{ES2}} F_B[|0_E\rangle\langle0_E| \otimes |\psi_{SA}\rangle\langle\psi_{SA}|, U_{ES1}^\dagger U_{ES2} \otimes I_A(|0_E\rangle\langle0_E| \otimes |\psi_{SA}\rangle\langle\psi_{SA}|)U_{ES2}^\dagger U_{ES1} \otimes I_A]$$

$$= \min_{|\psi_{SA}\rangle} \max_{U_{ES2}} |\langle\psi_{SA}|\langle0_E|U_{ES1}^\dagger U_{ES2} \otimes I_A|0_E\rangle|\psi_{SA}\rangle|$$

$$= \min_{|\psi_{SA}\rangle} \max_{U_{ES2}} |\langle\psi_{SA}|\langle0_E| \begin{bmatrix} K_{W_d} & * & * & \cdots & * \\ * & * & * & \cdots & * \\ \vdots & & & & \vdots \\ * & * & * & \cdots & * \end{bmatrix} \otimes I_A|0_E\rangle|\psi_{SA}\rangle| \quad (A8)$$

$$= \min_{|\psi_{SA}\rangle} \max_{W_d} |\langle\psi_{SA}|K_{W_d} \otimes I_A|\psi_{SA}\rangle|$$

$$= \min_{\rho^S} \max_{W_d} |Tr(\rho^S K_{W_d})|$$

$$= \min_{\rho^S} \max_{W_d} |Tr(\sum_{ij} w_{ij}\rho^S F_{1i}^\dagger F_{2j})|$$

$$= \min_{\rho^S} \max_{W_d} |\sum_{ij} w_{ij}Tr(\rho^S F_{1i}^\dagger F_{2j})|$$

$$= \min_{\rho^S} \max_{W_d} |Tr[W_d^T M(\rho^S)]|,$$

where $\rho^S = Tr_A(\rho^{SA})$ and $M(\rho^S)$ is a $d \times d$ matrix with its ij-entry equals to $Tr(\rho^S F_{1i}^\dagger F_{2j})$. As W_d is the first $d \times d$ submatrix of $W_p \in U(p)$, $\|W_d\| \leq 1$, here $\|\cdot\|$ denotes the operator norm which equals to the maximum singular value. Conversely any W_d satisfies $\|W_d\| \leq 1$ can be imbedded as a submatrix of a unitary[49], thus

$$\min_{|\psi_{SA}\rangle} F_B[(K_1 \otimes I_A)(|\psi_{SA}\rangle\langle\psi_{SA}|), (K_2 \otimes I_A)(|\psi_{SA}\rangle\langle\psi_{SA}|)] = \min_{\rho^S} \max_{\|W_d\| \leq 1} |Tr[W_d^T M(\rho^S)]|. \quad (A9)$$

Let $M(\rho^S) = UDV$ be the singular value decomposition of $M(\rho^S)$, where $U, V \in U(d)$ and $D = diag(s_1, s_2, \cdots, s_d)$, then $\max_{\|W_d\| \leq 1} |Tr[W_d^T M(\rho^S)]| = \|M(\rho^S)\|_1 = \sum_{i=1}^d s_i$, the equality is achieved with $W_d^T = V^\dagger U^\dagger$. This gives an analytical formula to compute the fidelity with any given pure input state

$$F_B[(K_1 \otimes I_A)(|\psi_{SA}\rangle\langle\psi_{SA}|), (K_2 \otimes I_A)(|\psi_{SA}\rangle\langle\psi_{SA}|))$$
$$= \|M(\rho^S)\|_1.$$

If we substitute K_1, K_2 with K_x and K_{x+dx}, we get $F_B(\rho_x, \rho_{x+dx}) = \|M(\rho^S)\|_1$, where $\rho_x = K_x \otimes I_A(|\psi_{SA}\rangle\langle\psi_{SA}|)$ and $\rho_{x+dx} = K_{x+dx} \otimes I_A(|\psi_{SA}\rangle\langle\psi_{SA}|)$, then using the connection between the Bures distance and the quantum Fisher information

$$d_{Bures}^2(\rho_x, \rho_{x+dx}) = \frac{1}{4}J(\rho_x)dx^2, \quad (A10)$$

we get the analytical formula for quantum Fisher information of the extended channel with any pure input state $|\psi_{SA}\rangle$, $J(K_x \otimes I_A(\rho^{SA})) = \lim_{dx\to0} \frac{8(1-\|M(\rho^S)\|_1)}{dx^2}$, where $\rho^S = Tr_A(|\psi_{SA}\rangle\langle\psi_{SA}|)$. Following the same line of argument, it can be shown that without auxillary system the same formula holds

$$F_B[K_x(|\psi_S\rangle\langle\psi_S|), K_{x+dx}(|\psi_S\rangle\langle\psi_S|)]$$
$$= \|M(\rho^S)\|_1, \quad (A11)$$

and $J(K_x(\rho^S)) = \lim_{dx\to0} \frac{8(1-\|M(\rho^S)\|_1)}{dx^2}$, just in this case $\rho^S = |\psi_S\rangle\langle\psi_S|$ is not the reduced state, but the pure input state $|\psi_S\rangle\langle\psi_S|$.

Note that when $W_d^T = V^\dagger U^\dagger$, $Tr[W_d^T M(\rho^S)] =$

$\|M(\rho^S)\|_1$ is a positive real number, thus

$$
\begin{aligned}
&\max_{\|W_d\|\le 1}|Tr[W_d^T M(\rho^S)]|\\
&=\max_{\|W_d\|\le 1}Re\{Tr[W_d^T M(\rho^S)]\}\\
&=\max_{\|W_d\|\le 1}\frac{1}{2}Tr\{W_d^T M(\rho^S)+[W_d^T M(\rho^S)]^\dagger\}\\
&=\max_{\|W_d\|\le 1}\frac{1}{2}Tr\{\rho^S K_{W_d}+[\rho^S K_{W_d}]^\dagger\}\\
&=\max_{\|W_d\|\le 1}\frac{1}{2}Tr[\rho^S K_{W_d}+K_{W_d}^\dagger\rho^S]\\
&=\max_{\|W_d\|\le 1}\frac{1}{2}Tr[\rho^S(K_{W_d}+K_{W_d}^\dagger)],
\end{aligned}
\tag{A12}
$$

then $\min_{|\psi_{SA}\rangle} F_B[(K_1 \otimes I_A)(|\psi_{SA}\rangle\langle\psi_{SA}|),(K_2 \otimes I_A)(|\psi_{SA}\rangle\langle\psi_{SA}|)] = \min_{\rho^S}\max_{\|W_d\|\le 1}\frac{1}{2}Tr[\rho^S(K_{W_d}+K_{W_d}^\dagger)]$, where $\rho^S = Tr_A(|\psi_{SA}\rangle\langle\psi_{SA}|)$ and $K_{W_d} = \sum_{i=1}^{d}\sum_{j=1}^{d}w_{ij}F_{1i}^\dagger F_{2j}$ with w_{ij} as the ij-th entry of W_d. For the extended channel since all the reduced states ρ^S forms a convex set, $\{W_d|\|W_d\|\le 1\}$ is convex and compact and $\frac{1}{2}Tr[\rho^S(K_{W_d}+K_{W_d}^\dagger)]$ is a bilinear function of entries of ρ^S and W_d, thus from Sion's theorem[50], the sequence of min-max can be exchanged, so

$$
\begin{aligned}
&\min_{\rho^S}\max_{\|W_d\|\le 1}\frac{1}{2}Tr[\rho^S(K_{W_d}+K_{W_d}^\dagger)]\\
&=\max_{\|W\|\le 1}\min_{\rho^S}\frac{1}{2}Tr[\rho^S(K_W+K_W^\dagger)]\\
&=\max_{\|W\|\le 1}\frac{1}{2}\lambda_{\min}(K_W+K_W^\dagger),
\end{aligned}
\tag{A13}
$$

here $\lambda_{\min}(K_W+K_W^\dagger)$ denotes the minimum eigenvalue of $K_W+K_W^\dagger$. We used a different symbol W after the exchange of min-max to emphasis that although the optimal value is not affected by the exchange of min-max, the

optimal point that achieves the optimal value can change. If we substitute K_1 with K_x and K_2 with K_{x+dx}, then we get the formula to compute the Bures angle between the extended channels

$$
\begin{aligned}
&\cos B(K_x\otimes I_A, K_{x+dx}\otimes I_A)\\
&=\min_{\rho^{SA}}F_B[(K_1\otimes I_A)(\rho^{SA}),(K_2\otimes I_A)(\rho^{SA})]\\
&=\max_{\|W\|\le 1}\frac{1}{2}\lambda_{\min}(K_W+K_W^\dagger),
\end{aligned}
\tag{A14}
$$

which then gives the maximal quantum Fisher information for the extended channel $K_x\otimes I_A$

$$
\begin{aligned}
&\max_{\rho^{SA}}J(K_x\otimes I_A(\rho^{SA}))\\
&=\lim_{dx\to 0}8\frac{1-\cos B(K_x\otimes I_A,K_{x+dx}\otimes I_A)}{dx^2}\\
&=\lim_{dx\to 0}\frac{8(1-\max_{\|W\|\le 1}\frac{1}{2}\lambda_{\min}(K_W+K_W^\dagger))}{dx^2},
\end{aligned}
\tag{A15}
$$

where $K_W=\sum_{ij}w_{ij}F_i^\dagger(x)F_j(x+dx)$.

2. Connection with previous studies

The above formula includes previous studies as a special case. Previous results as in [7, 8] state that for an extended channel $K_x\otimes I_A$ the maximal quantum Fisher information is given by

$$
\max J = 4\min_{\{\hat{F}_j(x)\}}\|\sum_{j=1}^{d}\dot{\hat{F}}_j^\dagger(x)\dot{\hat{F}}_j(x)\|
\tag{A16}
$$

where the minimization is over all smooth representations of equivalent Kraus operators of the channel K_x. Note that this can be equivalently written as

$$
\begin{aligned}
\max J &= 4\min_{\{\hat{F}_j(x)\}}\|\sum_{j=1}^{d}\dot{\hat{F}}_j^\dagger(x)\dot{\hat{F}}_j(x)\|\\
&= 4\min_{\{\hat{F}_j(x)\}}\|\sum_{j=1}^{d}\lim_{dx\to 0}\frac{(\hat{F}_j^\dagger(x+dx)-\hat{F}_j^\dagger(x))}{dx}\frac{(\hat{F}_j(x+dx)-\hat{F}_j(x))}{dx}\|\\
&= 4\min_{\{\hat{F}_j(x)\}}\|\frac{2I-\sum_{j=1}^{d}(\hat{F}_j^\dagger(x)\hat{F}_j(x+dx)+\hat{F}_j^\dagger(x+dx)\hat{F}_j(x))}{dx^2}\|\\
&= 4\frac{2-\max_{\{\hat{F}_j(x)\}}\lambda_{\min}[\sum_{j=1}^{d}(\hat{F}_j^\dagger(x)\hat{F}_j(x+dx)+\hat{F}_j^\dagger(x+dx)\hat{F}_j(x))]}{dx^2},
\end{aligned}
\tag{A17}
$$

where the optimization is over all smooth representations of equivalent Kraus operators. In previous studies the

equivalent Kraus operators are represented by $\hat{F}_j(x) =$

$\sum_{i=1}^{d}\omega_{ji}(x)F_i(x)$ and $\hat{F}_j(x+dx)=\sum_{i=1}^{d}\omega_{ji}(x+dx)F_i(x+dx)$, where $\omega_{ji}(x)$ is ji-entry of $W_E(x)\in U(d)$, and $W_E(x)$ is required to be smooth with respect to x. It is easy to see that in this case Eq.(A17) is just a special case of Eq.(A15) when W is restricted to taking the form $W_E^\dagger(x)W_E(x+dx)$.

We provide an example showing that the optimal W in Eq.(A15) can be non-unitary.

Consider two 8-dimensional channels with two Kraus operators each. The first channel $K_1(dx)$ has the following operators:

$$F_{11}=\mathrm{diag}(\alpha,\alpha,\alpha,\alpha,\alpha,\alpha,\alpha,\alpha) \tag{A18}$$
$$F_{12}=\mathrm{diag}(\beta,-\beta,\beta,-\beta,i\beta,-i\beta,i\beta,-i\beta) \tag{A19}$$

The second channel $K_2(dx)$ has the following:

$$F_{21}=F_{11} \tag{A20}$$
$$F_{22}=\mathrm{diag}(\beta,\beta,-\beta,-\beta,\beta,\beta,-\beta,-\beta) \tag{A21}$$

where $\alpha=\sqrt{1-dx^2}$ and $\beta=|dx|$ are real. Note that these two channels can be regarded as two neighbouring channels parameterized continuously in x, $\phi(x)=\begin{cases}K_1(x) & x\le 0\\ K_2(x) & x>0\end{cases}$.

We let $W=\begin{bmatrix}w_{11} & w_{12}\\ w_{21} & w_{22}\end{bmatrix}$, and compute $\frac{1}{2}\lambda_{\min}(K_W+K_W^\dagger)$, where $K_W=w_{11}F_{11}^\dagger F_{21}+w_{12}F_{11}^\dagger F_{22}+w_{21}F_{12}^\dagger F_{21}+w_{22}F_{12}^\dagger F_{22}$. First, note that if $W=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$, then $\frac{1}{2}\lambda_{\min}(K_W+K_W^\dagger)=\alpha^2$.

Now consider general W, since all matrices $F_{1i}^\dagger F_{2j}$ are diagonal in this case, we will only need to look at the diagonal elements. The diagonal elements of K_W are

$$\alpha(w_{11}\alpha+w_{12}\beta)+\beta(w_{21}\alpha+w_{22}\beta) \tag{A22}$$
$$+\quad -\quad +\tag{A23}$$
$$-\quad +\quad -\tag{A24}$$
$$-\quad -\quad -\tag{A25}$$
$$+\quad -i\quad +\tag{A26}$$
$$+\quad +i\quad +\tag{A27}$$
$$-\quad -i\quad -\tag{A28}$$
$$-\quad +i\quad -\tag{A29}$$

Thus, $Tr(K_W)=8\alpha^2 w_{11}$, thus $Tr(\frac{K_W+K_W^\dagger}{2})=8\alpha^2 Re(w_{11})$. If $Re(w_{11})<1$, the average eigenvalue of $\frac{K_W+K_W^\dagger}{2}$ will be smaller than α^2 thus the minimum eigenvalue will also be smaller than α^2, which is worse than the case of $W=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$, thus not optimal. So for optimal W, w_{11} must be 1. With $w_{11}=1$, the most general W with $\|W\|\le 1$ is $W=\begin{bmatrix}1 & 0\\ 0 & re^{i\theta}\end{bmatrix}$ with $0\le r\le 1$. The diagonal elements of K_W are then $\alpha^2\pm re^{i\theta}\beta^2$, $\alpha^2\pm ire^{i\theta}\beta^2$.

And the diagonal elements of $\frac{K_W+K_W^\dagger}{2}$ are

$$\alpha^2\pm\beta^2 r\cos(\theta), \alpha^2\pm\beta^2 r\sin(\theta).$$

If $r>0$, one of them must be less than α^2. So the optimal choice of W in this case is $W=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$, which is not unitary. And it is easy to check that the minimum value α^2 for $F_B[(K_1(dx)\otimes I_A)(\rho^{SA}),(K_2(dx)\otimes I_A)(\rho^{SA})]$ in this case can be achieved by preparing the input state as the maximally entangled state $\frac{1}{\sqrt{8}}\sum_{i=1}^{8}|ii\rangle$.

We also note that as the set $\{W|\|W\|\le 1\}$ is a convex set, this allowed a direct formulation with semi-definite programming as stated in the main text. While in previous studies W is restricted to be unitary which does not form a convex set. To circumvent the difficulty previous study has to resort to the Lie algebra of the unitaries and formulated the semi-definite programming there instead[10]. That, however, comes with a cost on the computational complexity, which can be seen by comparing the size of the constraining matrices in the semi-definite programming: the constraining matrices in the semi-definite programming of Eq.(13) have the total size of $2d+m_1$, while previous study needs a size of m_1+dm_2[10]. The difference can be significant when the system gets large(note that for generic channels d is in the order of $m_1 m_2$). For example for N-qubit system, $m_1=m_2=2^N$, the difference quickly becomes large with the increase of N.

3. Quantum Fisher information and optimal input states with dephasing noise and spontaneous emission

We give detailed calculation for the maximum quantum Fisher information and optimal input states for two examples, one with dephasing noise and the other with spontaneous emission.

Consider a channel with dephasing noise

$$K_x(\rho)=U(x)(\frac{1+\eta}{2}\rho+\frac{1-\eta}{2}\sigma_3\rho\sigma_3)U^\dagger(x),$$

where $U(x)=\exp(-i\frac{\sigma_3}{2}x),\eta\in[0,1]$. $\sigma_1=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$, $\sigma_2=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}$ and $\sigma_3=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$, $\eta\in[0,1]$. In this case $F_1(x)=\sqrt{\frac{1+\eta}{2}}U(x)$, $F_2(x)=\sqrt{\frac{1-\eta}{2}}\sigma_3 U(x)$. Let ρ^{SA} be a pure input state for the extended channel $K_x\otimes I_A$, and $\rho^S=Tr_A(\rho^{SA})$, then

$$M(\rho^S)$$
$$=\begin{pmatrix}Tr[\rho^S F_1^\dagger(x)F_1(x+dx)] & Tr[\rho^S F_1^\dagger(x)F_2(x+dx)]\\ Tr[\rho^S F_2^\dagger(x)F_1(x+dx)] & Tr[\rho^S F_2^\dagger(x)F_2(x+dx)]\end{pmatrix}$$
$$=\begin{pmatrix}\frac{1+\eta}{2}C & \frac{\sqrt{1-\eta^2}}{2}D\\ \frac{\sqrt{1-\eta^2}}{2}D & \frac{1-\eta}{2}C\end{pmatrix},$$

where $C = \rho_{11}^S \exp(-i\frac{dx}{2}) + \rho_{22}^S \exp(i\frac{dx}{2}), D = \rho_{11}^S \exp(-i\frac{dx}{2}) - \rho_{22}^S \exp(i\frac{dx}{2})$. Thus

$$M(\rho^S)M^\dagger(\rho^S) = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix},$$

where $m_{11} = \frac{(1+\eta)^2}{4}|C|^2 + \frac{(1-\eta^2)}{4}|D|^2$, $m_{12} = \frac{\sqrt{1-\eta^2}}{4}[(1+\eta)CD^* + (1-\eta)C^*D]$, $m_{21} = \frac{\sqrt{1-\eta^2}}{4}[(1+\eta)C^*D + (1-\eta)CD^*]$, $m_{22} = \frac{(1-\eta)^2}{4}|C|^2 + \frac{(1-\eta^2)}{4}|D|^2$. Denote s_1 and s_2 as singular values of $M(\rho^S)$, then $s_1^2 + s_2^2 = Tr(MM^\dagger)$ and $s_1 s_2 = |det(M)|$, thus

$$\begin{aligned}
&(s_1 + s_2)^2 \\
&= Tr(MM^\dagger) + 2|det(M)| \\
&= \frac{1+\eta^2}{2}|C|^2 + \frac{1-\eta^2}{2}|D|^2 + \frac{1-\eta^2}{2}|C^2 - D^2| \quad \text{(A30)} \\
&= (\rho_{11}^S + \rho_{22}^S)^2 - 2\rho_{11}^S\rho_{22}^S\eta^2[1 - \cos(dx)] \\
&= 1 - 4\rho_{11}^S\rho_{22}^S\eta^2 \sin^2 \frac{dx}{2}
\end{aligned}$$

To minimize $s_1 + s_2$, we just need to maximize $\rho_{11}^S\rho_{22}^S$, which is achieved with $\rho_{11}^S = \rho_{22}^S = \frac{1}{2}$. Thus

$$\cos[B(K_x \otimes I_A, K_{x+dx} \otimes I_A)] = \sqrt{1 - \eta^2 \sin^2 \frac{dx}{2}},$$

we then get the maximal quantum Fisher information $\lim_{dx \to 0} 8 \frac{1-\cos[B(K_x \otimes I_A, K_{x+dx} \otimes I_A)]}{dx^2} = \eta^2$, which is consistent with previous studies[13, 14]. In this case the optimal initial state is any pure state ρ^{SA} such that the reduced state ρ^S satisfies $\rho_{11}^S = \rho_{22}^S = \frac{1}{2}$. The simplest choice is the pure state $\frac{|0\rangle + |1\rangle}{\sqrt{2}}$, i.e. in this case ancillary system is not necessary, which means that the Bures angle for K_x and K_{x+dx} remains the same

$$\begin{aligned}
&\cos[B(K_x, K_{x+dx})] \\
&= \cos[B(K_x \otimes I_A, K_{x+dx} \otimes I_A)] \\
&= \sqrt{1 - \eta^2 \sin^2 \frac{dx}{2}},
\end{aligned} \quad \text{(A31)}$$

thus the maximal quantum Fisher information under the evolution K_x is also η^2. This is consistent with previous studies, however in our framework the optimal state comes out from the calculation naturally while in previous studies it depends on educated guess[13, 14].

For a channel with spontaneous emission, the Kraus operation takes form $K_x(\rho_0) = F_1(x)\rho_0 F_1^\dagger(x) + F_2(x)\rho_0 F_2^\dagger(x)$, where $F_1(x) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{\eta} \end{pmatrix} U(x)$, $F_2(x) = \begin{pmatrix} 0 & \sqrt{1-\eta} \\ 0 & 0 \end{pmatrix} U(x)$, $U(x) = \exp(-i\frac{\sigma_3}{2}x)$. Again let ρ^{SA} be a pure input state for the extended channel $K_x \otimes I_A$, and $\rho = Tr_A(\rho^{SA})$, then

$$M(\rho) = \begin{pmatrix} \rho_{11}e^{-i\frac{dx}{2}} + \eta\rho_{22}e^{i\frac{dx}{2}} & \sqrt{1-\eta}\rho_{21}e^{i(x+\frac{dx}{2})} \\ \sqrt{1-\eta}\rho_{12}e^{-i(x+\frac{dx}{2})} & (1-\eta)\rho_{22}e^{i\frac{dx}{2}} \end{pmatrix}.$$

Denote s_1 and s_2 as the singular values of $M(\rho)$ we can similarly get

$$\begin{aligned}
&(s_1 + s_2)^2 \\
&= \rho_{11}^2 + [\eta^2 + (1-\eta)^2]\rho_{22}^2 + 2\eta\rho_{11}\rho_{22}\cos(dx) \quad \text{(A32)} \\
&\quad + 2(1-\eta)(|\rho_{21}|^2 + |\rho_{11}\rho_{22} + \eta\rho_{22}^2 e^{idx} - |\rho_{21}|^2|).
\end{aligned}$$

First observe that

$$\begin{aligned}
&|\rho_{21}|^2 + |\rho_{11}\rho_{22} + \eta\rho_{22}^2 e^{idx} - |\rho_{21}|^2| \\
&\geq ||\rho_{21}|^2 + \rho_{11}\rho_{22} + \eta\rho_{22}^2 e^{idx} - |\rho_{21}|^2| \\
&= |\rho_{11}\rho_{22} + \eta\rho_{22}^2 e^{idx}| \\
&= \rho_{22}\sqrt{\rho_{11}^2 + 2\eta\rho_{11}\rho_{22}\cos dx + \eta^2\rho_{22}^2},
\end{aligned} \quad \text{(A33)}$$

where the equality is achieved with $\rho_{21} = 0$, so

$$\begin{aligned}
&(s_1 + s_2)^2 \\
&\geq \rho_{11}^2 + [\eta^2 + (1-\eta)^2]\rho_{22}^2 + 2\eta\rho_{11}\rho_{22}\cos(dx) \\
&\quad + 2(1-\eta)\rho_{22}\sqrt{\rho_{11}^2 + 2\eta\rho_{11}\rho_{22}\cos dx + \eta^2\rho_{22}^2} \\
&= [\sqrt{\rho_{11}^2 + 2\eta\rho_{11}\rho_{22}\cos dx + \eta^2\rho_{22}^2} + (1-\eta)\rho_{22}]^2 \\
&\geq [\sqrt{\frac{\eta + 2\eta\sqrt{\eta}\cos dx + \eta^2}{(1+\sqrt{\eta})^2}} + \frac{1-\eta}{1+\sqrt{\eta}}]^2.
\end{aligned} \quad \text{(A34)}$$

The last inequality is achieved by substituting $\rho_{11} = 1 - \rho_{22}$ into the equation and find the minimum of a single variable function, which is saturated with $\rho_{11} = \frac{\sqrt{\eta}}{1+\sqrt{\eta}}$ and $\rho_{22} = \frac{1}{1+\sqrt{\eta}}$. The minimum is thus achieved at $\rho = \begin{pmatrix} \frac{\sqrt{\eta}}{1+\sqrt{\eta}} & 0 \\ 0 & \frac{1}{1+\sqrt{\eta}} \end{pmatrix}$, which gives the Bures angle

$$\begin{aligned}
&\cos[B(K_x \otimes I_A, K_{x+dx} \otimes I_A)] \\
&= \sqrt{\frac{\eta + 2\eta\sqrt{\eta}\cos dx + \eta^2}{(1+\sqrt{\eta})^2} + \frac{1-\eta}{1+\sqrt{\eta}}},
\end{aligned} \quad \text{(A35)}$$

expanding both sides up to the second order of dx, we get the maximal quantum Fisher information $\lim_{dx \to 0} 4 \frac{B^2(K_x \otimes I_A, K_{x+dx} \otimes I_A)}{dx^2} = \frac{4\eta}{(1+\sqrt{\eta})^2}$. The optimal initial state in this case is any pure state ρ^{SA} with the reduced state equals to ρ, the simplest choice is $\sqrt{\frac{\sqrt{\eta}}{1+\sqrt{\eta}}}|00\rangle + \sqrt{\frac{1}{1+\sqrt{\eta}}}|11\rangle$.

The maximal quantum Fisher information for spontaneous emission without ancillary systems can also be calculated by imposing the state ρ used in $M(\rho)$ to be pure: $\rho = \begin{bmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{bmatrix}$ where a, b are complex numbers and $|a|^2 + |b|^2 = 1$. Eq. (A32) then becomes

$$(s_1 + s_2)^2 = |a|^4 + |b|^4 + 2[\eta\cos dx + (1-\eta)]|ab|^2 \quad \text{(A36)}$$

Taking the derivative with respect to $|a|^2$ gives the extremal point $|a|^2 = \frac{1}{2}$ corresponding to the minimum

value

$$(s_1 + s_2)^2 = 1 + \frac{\eta}{2}(\cos dx - 1) = 1 - \frac{1}{4}\eta dx^2 + O(dx^4).$$

(We know this is the minimum value by comparing to $|a|^2 = 0$.) This gives the Bures angle

$$\cos[B(K_x, K_{x+dx})] = 1 - \frac{1}{8}\eta dx^2 + O(dx^4). \quad (A37)$$

Finally, we get the maximal quantum Fisher information without using ancilla to be $\lim_{dx\to 0} 8\frac{1-\cos[B(K_x, K_{x+dx})]}{dx^2} = \eta$. The optimal input state is any pure state with $\rho_{11} = \rho_{22} = \frac{1}{2}$, which can be achieved by $\frac{|0\rangle + \exp(i\theta)|1\rangle}{\sqrt{2}}$ for any $\theta \in \mathbb{R}$.

[1] V. Giovannetti, S. Lloyd, & L. Maccone, *Nature Photonics.* **5**, 222 (2011).
[2] D. J. Wineland, J. J. Bollinger, W. M. Itano, F.L. Moore, & D. J. Heinzen, *Phys. Rev. A* **46**, R6797-R6800 (1992).
[3] C. M. Caves, *Phys. Rev. D* **23**, 1693-1708 (1981).
[4] H. Lee, P. Kok, & J.P. Dowling, *J. Mod. Opt.* **49**, 2325-2338 (2002).
[5] S. L. Braunstein, *Phys. Rev. Lett.* **69**, 3598 (1992).
[6] V. Giovannetti, S. Lloyd, & L. Maccone, *Science* **306**, 1330-1336 (2004).
[7] A. Fujiwara,& H. Imai, *J. Phys. A: Math. Theor.* **41**, 255304 (2008).
[8] B.M. Escher, R.L. de Matos Filho, & L. Davidovich, *Nature Phys.* **7**, 406 (2011).
[9] M. Tsang, *New J. Phys.* **15**, 073005 (2013)
[10] R. Demkowicz-Dobrzański, J. Kołodyński, & M. Guta, *Nature Comm.* **3**, 1063 (2012).
[11] S. Knysh, V. N. Smelyanskiy, & G. A. Durkin, *Phys. Rev. A.* **83**, 021804 (2011).
[12] S. Knysh, E. Chen, & G. Durkin, arXiv:1402.0495 (2014).
[13] J. Kołodyński, & R. Demkowicz-Dobrzański, *New Journal of Physics* **15**, 073043 (2013).
[14] R. Demkowicz-Dobrzański, & L. Maccone, *Phys. Rev. Lett.* **113**, 250801 (2014).
[15] S. Alipour, M. Mehboudi, & A.T. Rezakhani, *Phys. Rev. Lett.* **112**, 120405 (2014).
[16] A.W. Chin, S. F. Huelga, and M. B. Plenio, *Phys. Rev. Lett.* **109**, 233601 (2012).
[17] M. Tsang, H. M. Wiseman, and C. M. Caves Phys. Rev. Lett. 106, 090401 (2011).
[18] D. W. Berry, M. J. W. Hall, and H. M. Wiseman Phys. Rev. Lett. 111, 113601(2013).
[19] D. W. Berry, M. Tsang, M. J.W. Hall, and H. M. Wiseman, Phys. Rev. X 5, 031018 (2015).
[20] H. Cramér, Mathematical Methods of Statistics. (Princeton University, Princeton NJ, 1946).
[21] Rao, C. R. *Bull. Calcutta Math. Soc.* **37**, 81 (1945).
[22] C. W. Helstrom, Quantum Detection and Estimation Theory. (Academic Press, New York, 1976).
[23] A. S. Holevo, Probabilistic and Statistical Aspect of Quantum Theory. (North-Holland, Amsterdam, 1982).
[24] R. A. Fisher, *Proc. Cambr. Phil. Soc.* **22**, 700 (1925).
[25] S. L. Braunstein, & C. M. Caves, *Phys. Rev. Lett.* **72**, 3439 (1994).
[26] S. L. Braunstein, C. M. Caves, & G. J. Milburn, *Annals of Physics* **247**, 135-173 (1996).
[27] V. Giovannetti, S. Lloyd, & L. Maccone, *Phys. Rev. Lett.* **96**, 010401 (2006).
[28] B. M. Escher, L. Davidovich, N. Zagury, & R. L. de Matos Filho, *Phys. Rev. Lett.* **109**, 190404 (2012).
[29] A. Uhlmann, Quantum Groups and Related Topics: Proceedings of the First Max Born Symposium, Gielerak, R.; Lukierski, J.; Popowicz, Z. (Eds.), pp 267 (1992).
[30] W.K. Wootters, *Phys. Rev. D*, **23**, 357 (1981).
[31] H. F. Chau, *Quant. Inf. Compu.* **11**, 0721 (2011).
[32] C.-H. F. Fung, & H. F. Chau, *Phys. Rev. A* **90**, 022333 (2014).
[33] C.-H. F. Fung, H. F. Chau, C.K. Li, & N.S. Sze, Quantum Information and Computation, Vol. 15, No. 7&8, 0685-0693 (2015).
[34] H.D. Yuan, & C.-H. F. Fung, arXiv: 1506.00819 (2015).
[35] H.D. Yuan, & C.-H. F. Fung, Phys. Rev. Lett. 115, 110401 (2015)
[36] D. J. C. Bures, *Trans. Am. Math. Soc.*, **135**, 199 (1969).
[37] J. Kołodyński, arXiv:1409.0535 (2014).
[38] Please see appendix for detail.
[39] A. Childs, J. Preskill, and J. Renes. Quantum information and precision measurement. *Journal of Modern Optics*, 47(2–3):155–176, 2000.
[40] A. Acín. Statistical distinguishability between unitary operations. *Physical Review Letters*, 87(17):177901, 2001.
[41] R. Y. Duan, Y. Feng, and M. S. Ying, "Entanglement is Not Necessary for Perfect Discrimination between Unitary Operations", Phys. Rev. Lett. 98, 100503 (2007).
[42] D. W. Berry, H. M. Wiseman, *Phys. Rev. Lett.* 85, 5098 (2000).
[43] R. Demkowicz-Dobrzanski, U. Dorner, B. J. Smith, J. S. Lundeen, W. Wasilewski, K. Banaszek, and I. A. Walmsley, *Phys. Rev. A* 80, 013825 (2009).
[44] R. Nair, B.J. Yen, *Phys. Rev. Lett.* 107, 193602 (2011).
[45] F. Fröwis, M. Skotiniotis, B. Kraus, and W. Dür, New Journal of Physics 16, 083010 (2014).
[46] Fuel,M., Hindi,H., & Boyd,S.P., Proceedings of the American Control Conference, Arlington, VA June 25-27(2001).
[47] M. Grant, & S. Boyd, CVX: Matlab software for disciplined convex programming, http://cvxr.com/cvx/ (2011).
[48] A. Uhlmann, *Rep. Math. Phys.* **9**, 273-279 (1976).
[49] Choi, M.D., & Li, C.K., *J. Operator Theory.* **46**, 435 (2001).
[50] Sion, M., *Pac. J. Math.* **8**, 171-176 (1958).